DOCUMENT RESUME

ED 151 686 CG 012 292

AUTHOR Underwood, Benton J.; Malmi, Robert A.
TITLE Recall and Recognition of Tasks Learned Simultaneously.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE Jul 77
CONTRACT N00014-76-C-0270
NOTE 29p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Adults; *Associative Learning; *Learning; *Mnemonics; *Recall (Psychological); Research Projects; *Retention; *Transfer of Training

ABSTRACT

Three clearly distinguishable lists were given simultaneously for learning. In one condition, subjects recalled all three lists; in a second condition, each list was given a different type of retention test although one of the lists was recalled. All subjects were fully informed of the materials and tests. The critical interest was in the recall of the same list in the two conditions. Several lines of thought led to the expectation that recall would be better when three different retention tests were used than when all lists were recalled. Three experiments gave little support to the expectation. To enhance differential encoding of the lists to match the retention tests, a fourth experiment was conducted that included a condition in which subjects learned each list separately before simultaneous learning. No evidence for differential encoding was found. There was relatively little transfer from single-list to simultaneous learning of three lists represents a marked change in context. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED151688

RECALL AND RECOGNITION OF TASKS LEARNED SIMULTANEOUSLY

Benton J. Underwood and Robert A. Malmi

Northwestern University

July, 1977

Project RR 042-06
Sponsored by
Personnel & Training Research Programs
Psychological Sciences Division
Office of Naval Research
Arlington, Virginia
Contract No. N00014-76-C-0270

Approved for public release; distribution unlimited.
Reproduction in whole or in part is permitted
for any purpose of the United States Government.

CG012292

Recall and Recognition of Tasks Learned Simultaneously

Benton J. Underwood and Robert A. Malmi

Northwestern University

Suppose we have two clearly distinguishable sets of words, A and B, which are to be presented as free-recall lists. Rather than presenting each set separately for learning, we present them simultaneously. On the study phase, a word from Set A and a word from Set B are shown together, then another pair is shown, and so on, until all words in both sets have been presented. The subjects are fully informed in that they are told at the beginning of the study phase that there are two sets of words, and are told further just how one set is distinguishable from the other. One group is told that on the memory test the A words would be recalled first, following which the B words would be recalled. Another group of subjects is told that they would be first asked to recall the A words, and then would be given a recognition test of memory for the B words. We will identify the first group as Group RR (recall-recall), and the second as Group RD (recall-recognition, where D represents discrimination, or d'). Our critical interest is in the recall of the A words. What can be predicted with respect to the recall of the A words for the two groups? We will identify two classes or types of theory which would seem to predict that the recall of the A words would be higher for Group RD than for Group RR.

Theories which might fall in the first class would predict the better recall of A words by the subjects in Group RD than by the subjects in Group RR because preparing for two different retention tests allows the subject to control his learning efforts to meet the demands of the task. One such possibility will be described. Investigators generally accept the idea that

for the usual experiment, recognition will be better than recall, although just how tests of recall and recognition can be made equivalent is not apparent. Still, the generalization is probably correct in the sense that we may often recognize items that we cannot recall, whereas the reverse possibility occurs only under unusual circumstances. It is probably also correct to say that the usual college student knows that to learn to recognize an item takes less study time than to learn to recall it. These premises suggest that under simultaneous learning the subjects in Group RD might allocate their study time differentially between the two lists. If each presentation of two words (one from each set) is for 6 seconds, the subject might spend 4 second studying the item to be recalled, and 2 seconds studying the item to appear on the recognition test. The subjects in Group RR should allocate an equal amount of time to each item because both sets are to be recalled. The prediction is clear, namely, that the recall of the A items will be higher for Group RD than for Group RR.

The same prediction can be reached by several routes which initially assume differential encoding for recall and for recognition. Differential encoding means that the encoded words for recall consist of different types of information from the types constituting the memories for words encoded for recognition. A second assumption would provide for some degree of independence in storage "space" for the two sets of words given differential encodings, and also make the total space available greater when two types of encodings are used rather than one. A third assumption must provide a decreasing encoding or storage function per unit of study time, e.g., the greater the space already occupied, the more difficult it is to fill the re-

maining space. Assumptions such as these can lead to the prediction that recall of the A words would be better for Group RD than for Group RR.

As is apparent, the first class of theories is based on the control of study time by the subject, and falls back on the simple fact that the amount learned is directly related to study time. Theories of this type might be viewed as being rather uninteresting as compared with theories of the second type which emphasize differential encoding. Nevertheless, if the outcome is as predicted by both classes of theories, further steps will have to be taken to determine which seems to best account for the evidence.

The failure to find a difference in the recall of the A words for the two groups would, of course, call into question the assumptions underlying the theories. Perhaps of greatest theoretical interest would be the apparent denial of the differential encoding assumption. It would deny the possibility that when two tasks are given for simultaneous learning (one to be tested by recall, one by recognition), the subject deliberately encodes them differently. Insofar as we have been able to discover, the experimental literature does not contain evidence which is directly relevant to the empirical question we are asking. In fact, we have not found studies in which two tasks, normally learned separately, are learned simultaneously but tested as two different tasks. There have been, of course, many studies in which a second task has been presented along with the target task with the intent of gauging the difficulty of various target tasks in terms of variation in the performance on the second or auxiliary tasks. Work in this and related areas is reviewed by Kahneman (1973), Kerr (1973), Garner (1974), and Norman and Bobrow (1975).

Experiment 1

This experiment consisted of three conditions which will be identified as RRR, RRD, and RFD. In the illustrations used above, subjects were said to have been given two lists to learn simultaneously. In all of the studies to be reported, three lists were used along with three different types of retention tests. The lists will be identified as 1, 2, and 3 as determined by the order in which the lists were tested. Under Condition RRR, the subjects recalled all three lists in order; under Condition RRD, Lists 1 and 2 were recalled in order, followed by a recognition test for List 3. Under Condition RFD, the subjects recalled List 1, were given a frequency-judging test (F) on List 2, and a recognition test on List 3. The critical interest centers on the recall of List 1 as a function of the number (1, 2, 3) of different types of retention tests. A further examination of the influence of number of different retention tests can be made by comparing the recall of List 2 for conditions RRR and RRD, where the number of different types of retention tests are one and two, respectively. Finally, it can be seen that by comparing recognition scores on List 3 for Conditions RRD and RFD, we can determine if recognition differs as a function of number of different types of retention tests (two versus three).

Method

Lists. Three classes of materials were made into lists, each of the three lists consisting of 27 pairs of words. The intent was to make each class clearly distinguishable from the others. We used the Battig-Montague (1969) norms to construct 27 pairs made up of animal names on the left, girls' first names on the right, e.g., Buffalo-Donna, and this was called

List 1. List 2 consisted of pairs in which names of cities were on the left, names of birds on the right, e.g., Berlin-Ostrich. The third class (List 3) consisted of names of colleges or universities on the left, boys' first names on the right, e.g., Yale-Harvey. In each of the lists, 9 pairs were designated randomly to be presented once, 9 pairs to be presented twice, and 9 pairs to be presented three times. Then, the three different lists were brought together into a single list consisting of 54 sets, each set containing three pairs, one from each class. Tuus, the three pairs in the first set consisted of Columbia-Henry, Buffalo-Donna, Rome-Thrush, horizontally displayed through the window of the memory drum. The second set consisted of Seattle-Sparrow, Millikin-Brian, Donkey-Louise. The 162 positions (54 sets of three pairs each) contained 9 pairs at frequencies 1, 2, and 3 for each class as described earlier. The position of the three pairs within each set of three was random.

Conditions. All subjects were presented the same study list. The three conditions differed in terms of the nature of the memory tests given for the three lists. All subjects were tested in the same order on the three lists. For Condition RRR, the subjects recalled all three lists in order. For Condition RRD, the subjects recalled List 1, then recalled List 2, and then were given a recognition test (D) on List 3. On this recognition test, 9 new pairs were randomized among the 27 old pairs to produce a 36 pair YES-NO recognition test. For the third condition (RFD), the subjects recalled List 1, gave frequency judgments for List 2, and then were given a recognition test on List 3. For the frequency-judgment test for List 2, 9 new pairs were added to the 27 appearing in the study list, and

the subjects circled a number for each pair (0, 1, 2, 3) to indicate their judgments of the frequency with which the pair had been presented.

Procedure and subjects. Thirty subjects were assigned to each of the three conditions by a block-randomized schedule of conditions. The subjects were fully instructed as to the nature of the list and the nature of the memory tests appropriate for each condition. After the initial instructions, the list was presented for the first study trial at a 12-second rate, i.e., 12 seconds for each set containing three pairs. The list included one primacy set and one recency set. On the recall tests the subject was presented a prepared sheet on which the nature of the pairs to be recalled was described, and for which the appropriate pairs from the primacy and recency sets were used as illustrations. The subjects were given as much recall time as needed or desired, and they were urged by the experementer to "dig a little deeper" when it appeared that recall attempts were about to be terminated by the subject. After the subject gave up on his attempts to recall pairs from List 1, the experimenter moved to List 2 which was handled in the same manner as List 1 for Conditions RRR and RRD. For Condition RFD, List 2 was given as a frequency-judging test. The instructions were printed on the test sheet and these were read by the experimenter as the subject followed. The subjects were required to make a decision on all 36 pairs, the pairs being listed in alphabetical order.

The subjects in Condition RRR recalled List 3, and the procedure was exactly the same as for Lists 1 and 2. For Conditions RFD and RRD, the memory for the third list was tested by recognition. The instructions were printed on the test sheet and again the experimenter read these instructions

as the subject followed. The 36 pairs were alphabetized on the test sheet and the subjects circled YES or NO to indicate their decisions for each pair.

After the first set of tests had been completed, the subjects were given a second study trial followed by a second set of tests exactly as had been given after the first study trial. The order of the sets on the two study trials was exactly the same.

Results

Recall and recognition. The recall values of interest were those for List 1 for all three conditions, and for List 2 under conditions RRR and RRD. The mean recall scores for these two lists for each trial are shown in Figure 1. For both conditions, recall increased as the number of lists to be recalled decreased. We quickly add, however, that the statistical support for this conclusion is dubious. The differences among the three conditions for List 1 were not reliable, F (2,87) = 2.96, $p > .05$. For List 2, the F for conditions was less than one. But, for both lists, the interaction between trials and conditions attained the .05 level of significance, indicating that the differences among conditions were increasing as trials continued. It will be remembered that the pairs were presented with different frequencies, nine pairs each being given frequencies of 1, 2, and 3. Recall, of course, increased directly as frequency increased. An evaluation of condition differences was made for the recall (both trials combined) of the nine pairs presented once on the study trial. The means were 2.83, 4.00, and 4.90 for Conditions RRR, RRD, and RFD, respectively, and the differences would be judged reliable, F (2,87) = 5.21, $p < .01$. Still, when all

frequency levels were included in the analysis, the interaction between conditions and frequency levels was not reliable. It appears to us that these data indicate that if there is an influence on recall which is associated with the number of different retention tests, its magnitude is small.

If disproportionate effort is allocated to recall tasks during study, the recognition scores for Condition RRD should be poorer than those for Condition RFD. In fact, the mean sum of the misses and false alarms (combined for both trials) was actually greater under Condition RRD (9.60) than under Condition RRD (7.63), although the difference was not reliable ($t = 1.16$).

Frequency judgments. These judgments were given to the pairs of List 2 under Condition RFD. Our comments will be brief because these scores do not speak to the basic question. The mean correlation between the true frequencies and judged frequencies for the 36 pairs was .79 on the first trial, .78 on the second. The failure of the mean correlation to increase from Trial 1 to Trial 2 has been found in other experiments as yet unpublished. The reason for the lack of increase is unknown. The average judged frequencies for both trials combined were .24, .97, 1.60, and 2.19 for frequencies 0, 1, 2, and 3, respectively.

Correlations. The purpose of examining intertask correlations is to make inferences about encoding processes. High intertask correlations do not tell us much about the independence of encoding processes, but correlations not differing statistically from zero do. When the performance scores on two tasks are unrelated, it must mean that the processes underlying the two tasks are independent. One of the issues of interest in the present studies concerns the possibility of subjects encoding recall and recogni-

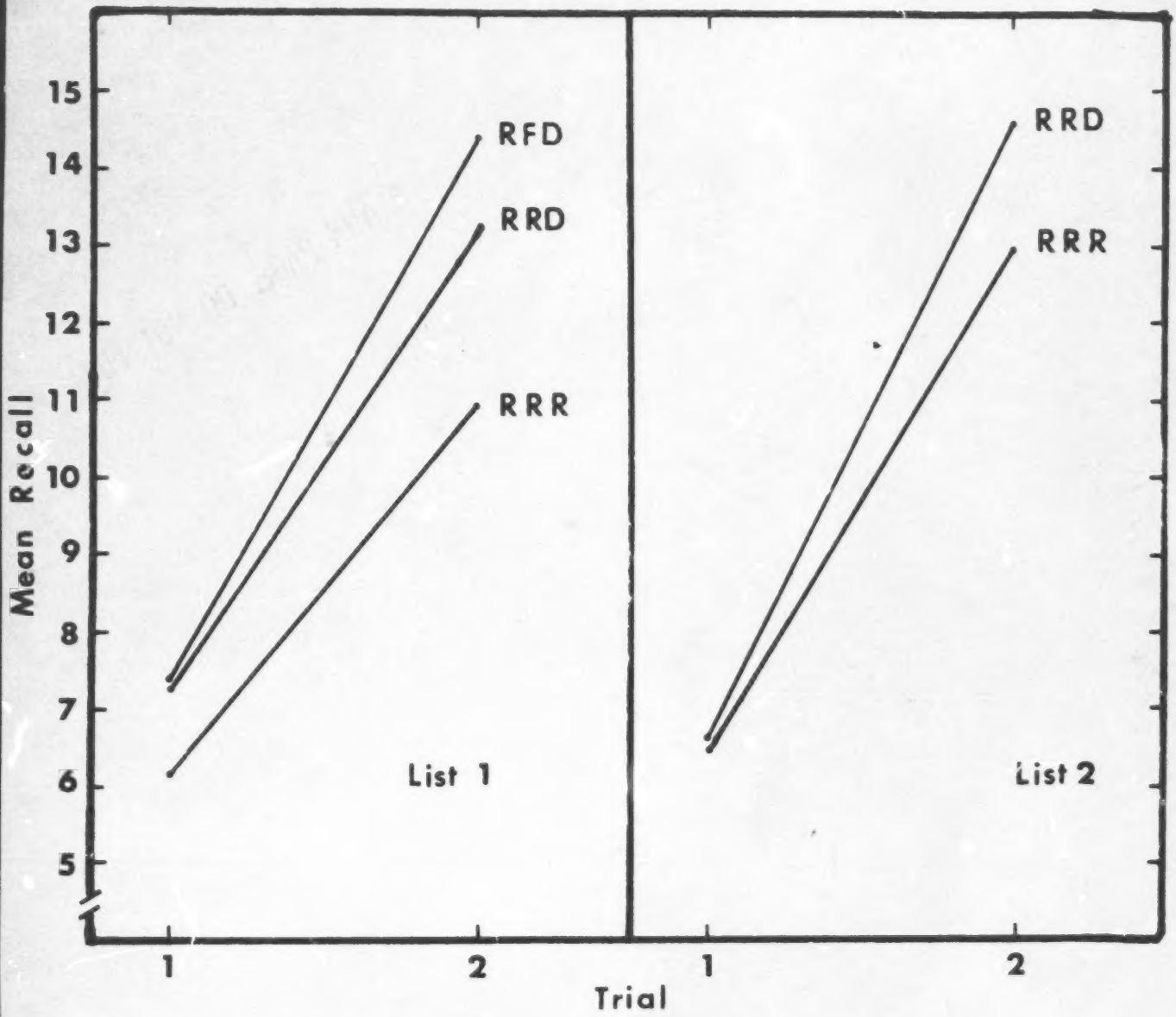

Figure 1. Recall as a function of list, condition, and trial (Experiment 1).

tion tasks differently. A zero correlation between the scores on the retention tests for recall and recognition would support the idea of differential encoding. To expect to interpret zero correlations in a meaningful way requires that the scores be reliable. For the present studies we have used the correlations between the scores on trial 1 and trial 2 of a given task to estimate reliability.

There were six lists that were recalled. The intertrial correlations varied between .77 and .88. For recognition, the reliabilities were .62 and .79 for Conditions RRD and RFD in order, and the reliability of the frequency judgments for Condition RFD was .86.

The intertask correlations are shown in Table 1. The most noteworthy fact implied by these correlations is the rather complete lack of discrimination associated with types of retention tests. The three correlations for Condition RRR represent the relationships among three recall tests. The three correlations for Condition RFD represent the correlations among three different types of retention tests. Given the comparability in the relationships, one would be hard pressed, to say the least, to make a case for the proposition that different processes underlie the three different retention tests.

Discussion

The results gave weak evidence for a relationship between recall and number of different types of retention tests expected. Furthermore, recognition was uninfluenced by the number of different types of retention tests.

The correlational evidence was most surprising in that under Condition RFD the interlist correlations were quite substantial and numerically averaged higher than for Condition RRR where the subject recalled all three lists. If subjects in Condition RFD were encoding the three tasks differentially, the skills at doing so were apparently intercorrelated. On the other hand, the subjects may have simply encoded all pairs the same.

The subjects appeared to have no problem distinguishing among the classes of the material in the three lists. They never recalled a pair from a wrong list. Still, the distinction among the three lists may have been more apparent to the experimenters than to the subjects; the subjects may have made little effort to classify the items because to do so was not easy and for some subjects, perhaps, a waste of good study time. We decided, therefore, to construct a new set of materials which could be more readily classified into three domains.

Experiment 2

The three different types of pairs presented the subjects in Experiment 1 seemed to be clearly distinguishable from each other. Still, this may not have been true when subjects attempted to determine quickly the nature of the retention test to be given for a particular item. The distinction among the pairs would have to be made on the basis of semantic differences, and it is possible that some negative effects in encoding may have been present in Condition RFD as the subjects attempted to classify each item prior to differential encoding. In Experiment 2 we established differences among the three types of items on both semantic and orthographic dimensions. It thus became possible for the subjects to determine how an item would be

Table 1

Interlist Correlations Based on Scores Across Two Trials for Experiment 1

Condition	List 1 & List 2	List 1 & List 3	List 2 & List 3
RRR	.65	.62	.31
RRD	.61	.51	.29
RFD	.56	.63	.66

NOTE: With 28 df, a correlation of .36 is needed for $p = .05$, .46 for $p = .01$.

tested on the basis of a primitive analysis using only orthographic cues. Classification decisions could have been made very rapidly so that if differential encoding was to occur, almost the entire study time could be devoted to it. We used only two conditions for this experiment, RRR and RFD. The procedures were almost exactly the same as those used in Experiment 1, the essential differences being in the materials used for constructing the three lists.

Method

Lists. List 1 consisted of 24 pairs of names, with boys' names on the left, girls' names on the right. The first letter was capitalized, e.g., Howard-Rita. Each pair occurred twice in the study list, once in each half. List 2 consisted of 24 animal names which were all printed in capital letters, e.g., MOUSE. Eight words were assigned to each frequency (1, 2, 3). Under Condition RRR, of course, the subject recalled as many of the 24 animal names as possible. Under Condition RFD, the 24 words were mixed with eight new words and the subject made absolute frequency judgments for the 32 words. List 3 consisted of two-word phrases, the two words being separated by a hyphen. These phrases were printed in lower case on the memory-drum tape. The phrases were all common sequences such as foul-line, picture-frame, and and barbed-wire. Each was presented once on the study trial. These pairs were recalled by the subjects in Condition RRR, and for the subjects in Condition RFD, the 24 phrases were mixed with eight new phrases and presented as a YES-NO recognition test.

These three lists were merged, of course, to form a single study list in which the three sublists were learned simultaneously. To accommodate the frequencies, 40 sets of three stimuli each were required. No primacy

or recency buffers were included. Within an exposure set, all of the three items were from the same sublist, i.e., they were all animal names, or all name pairs, or all common phrases. If subjects encoded the lists differently, the arrangement allowed all items being exposed at the moment to be encoded in the same way. Also, the method of presentation would make it difficult for a subject to allot a disproportionate amount of time to the list to be recalled. To do so would require that pairs previously presented be recalled and rehearsed while the drum was exhibiting items for recognition and for frequency judgments.

Procedures and subjects. All subjects were fully informed of the nature of the materials and how they would be tested. All details were the same as for Experiment 2. Again, 30 subjects were assigned to each condition by a block-randomized schedule.

Results

Recall. The critical comparison involves the recall of List 1 (pairs of names) under the two conditions. For Condition RRR, the means and standard deviations (in parenthesis) on the two trials were 7.27 (3.22) and 13.70 (4.64). For Condition RFD, the corresponding values were 6.53 (3.64) and 12.43 (4.72). Thus the subjects who were required to recall all tasks actually did a little better on both trials than did those who recalled only the first list, although, of course the differences were not reliable ($F < 1$). The subjects in Condition RRR also recalled the other two lists. The recall of List 2 (made up of animal names) was high, the means for the two trials being 13.97 (3.50) and 18.97 (3.25). The means for List 3 (two-word phrases) were 4.63 (2.19) and 11.70 (4.36) for the two trials.

Recognition and frequency judgments. For Condition RFD, the mean correlation between true and judged frequency for the items in List 2 was .81 for the first trial, .83 for the second. For recognition on the third list, the mean sum of the misses and false alarms was 6.53 (4.67) and 3.60 (3.34) for the two trials in order.

Correlations. The intertrial correlations for all lists were quite high, varying between .65 and .82. The interlist correlations led to much the same conclusion as had been reached for Experiment 1 (Table 1). For Condition RRR, the three correlations (1 x 2; 1 x 3; 2 x 3) based on scores summed over trials were .40, .34, and .67. For Condition RFD, where the correlations represent the relationships between scores on different types of memory tests, the values were .42, .38, and .42. That the correlations for both lists were somewhat lower than those found for Experiment 1 probably reflects the fact that the lists used in the present experiment were more heterogeneous (on several dimensions) than those of Experiment 1.

Discussion

The results of the present experiment were unequivocal; recall was not related to the number of different retention tasks expected. If the lists were encoded differentially to match the different retention tests, there was no evidence of this in recall differences or in differences among the correlations. It will be remembered that in the introduction we noted that two lines of thought would lead to the expectation that recall of the first list would be higher for Condition RFD than for Condition RRR. One of these was based on differential encoding, a possibility which, in view of our results, was beginning to seem unlikely. The other was based on the some-

what uninteresting idea that the subject might allocate more of the available study time to items to be recalled than to items which would be tested for recognition. As discussed earlier, because a given exposure set in the present experiment consisted of items from the same list, it would have been difficult for the subject to divide the study time unequally. Therefore, we did the third experiment to allow the subject this possibility, just as it was possible in Experiment 1.

Experiment 3

Method

This study was conducted in the same way as Experiment 2; only a slight change was made in the presentation of the materials. In Experiment 2, the three stimuli presented together during the 12-second exposure period were all of the same class. In Experiment 3, we simply arranged the items so that on a given 12-second exposure each of the three classes or lists was represented by one item. Thus, the presentation was comparable to that used in Experiment 1. The 30 subjects in each condition were assigned by a block-randomized schedule.

Results

Recall. The recall scores for List 1 were comparable to those found in Experiment 2. For Condition RRR, the mean values were 6.97 (4.40) and 12.53 (5.00) for the two trials. The corresponding values for Condition RFD were 7.30 (3.87) and 13.77 (4.68). The direction of the small difference was opposite to that reported for Experiment 2. The smallness of the differences for both experiments must be emphasized. We did an analysis of variance for recall of List 1 for the two experiments for the two conditions to see if the interaction between conditions and experiments was reliable.

It was not (F = 1.36).

For trials 1 and 2 of List 2, the mean recall was 13.00 (3.04) and 17.63 (3.21). For List 3 the values were 3.83 (1.72) and 9.67 (3.46). All of these values are a little lower than those for Experiment 2.

Recognition and frequency judgments. The mean correlation between true and judged frequency for List 2 under Condition RFD was .76 for the first trial and .77 for the second. The sums of the misses and false alarms for List 3 were 8.03 (3.10) and 5.97 (2.93) for the two trials in order. None of these findings is related to the central issue under investigation, but they do show that the subjects learned all lists to some degree.

Correlations. The reliabilities, as shown by intertrial correlations, were a little lower and a little more varied than those observed for Experiment 2. For Condition RRR, the three correlations in order (lists 1, 2, 3) were .92, .48, and .58. The values for Condition FRD were .82, .62, and .55. For the intertask correlations, some changes were also observed. The three correlations for Condition RRR in order (1 x 2, 1 x 3, 2 x 3) were .61, .61, and .67. For Condition RFD, the values were -.01, -.09, and .27. The values for this latter condition for Experiment 2 were .42, .38, and .42. However, none of the differences between the correlations in the two experiments for Condition RFD was found to be statistically reliable. Furthermore, as we will see, the correlations were not well replicated in Experiment 4.

Discussion

The data for the three experiments as a whole seem to indicate that we have nothing to explain, but in so saying, we are denying the assumptions of the two classes of theories discussed in the introduction. The results seem fairly definitive with regard to the first type of theory, the type

which proposed that the subject allocated his study time differentially for recall and for recognition. In Experiment 2, three different items from the same list were exposed together; in Experiment 3 the three items consisted of one from each list. It is our belief that the subject could more readily devote differential study time to items in the lists in Experiment 3 than in Experiment 2, but if this did happen there was no clear evidence that the recall of the first list was changed thereby.

The second class of theories depends upon an assumption of differential encoding for recall and recognition tests. We have found little evidence to support this idea. Still, our situation may not be a good one in which to encourage differential encoding. The subjects may be so busy merely trying to learn something that it would be expecting too much to assume that they could carry out the classification of the items and subsequent differential encoding. With this possibility as background, we attempted in Experiment 4 to make it easy for the subject to encode differentially.

Experiment 4

If differential encoding of tasks does take place, it should most readily occur for tasks learned singly. Thus, if subjects are given a single list under recall instructions, and at another time a second list under recognition instructions, differential encoding should occur if it is going to occur at all. Therefore, we attempted to induce our subjects to encode differentially by giving them an initial trial on each list alone. Following this, the subjects were transferred to simultaneous learning of the three tasks just as in the previous experiments. This group will be designated as representing Condition T-RFD, where the T indicates the initial

study-test trial on each list independently. The second group represented Condition RFD, and the procedure for this condition was exactly the same as in Experiment 3. If the initial study-test trial on each task alone for Condition T-RFD produces differential encoding (at least for recall and recognition), and if this carries over to simultaneous learning, then, as expected by theories emphasizing differential encoding, recall performance on the first test trial of simultaneous learning under Condition T-RFD should be greater than recall performance on the second test trial for Condition RFD.

Method

The materials and procedures were exactly the same as for Experiment 3 as far as simultaneous learning was concerned. The difference in the two experiments was represented by the initial trial on each task alone for the subjects in Condition T-RFD. The subjects were first given the 24 name pairs in the same order as would occur in the subsequent simultaneous learning. Each pair occurred twice, and the exposure period was 4 seconds. We simply allowed one-third of the exposure time given in simultaneous learning where three tasks were being learned. After the single study trial, recall was taken. Then, the second list was given for a study trial (the order and frequency being the same as for simultaneous learning) at a 4-second rate, followed by the frequency-judging test. Finally, the 24 pairs for recognition were given a study and test trial. Immediately thereafter, the subjects were given the simultaneous list (12-second rate) for two study and test trials to match the two trials given to the subjects in Condition RFD.

It should be emphasized that the subjects in Condition T-RFD were fully

informed throughout. Before any list was given for study, the subjects were told that they would be given three different lists alone. The materials in each list were described as well as the nature of the retention tests. The subjects knew that after having the trials on each list alone, they would be given further learning trials on the lists combined for simultaneous presentation.

Thirty subjects were assigned to each of the two groups by a block-randomized schedule.

Results

Recall. The mean numbers of correct responses on each trial under each condition are shown in Figure 2. We will speak of three trials for Condition T-RFD where the first trial indicates the performance on the tasks when presented singly, and trials 2 and 3 represent simultaneous learning. It is evident from Figure 2 that recall for trial 2 under Condition T-RFD was not superior to the recall on the second trial for Condition RFD as might be expected if differential encoding occurred. In fact, the increase in performance between trial 1 and trial 2 is less under Condition T-RFD than under Condition RFD as shown by a significant interaction, F (1,58) = 5.34, $p < .05$. Overall, there was a small positive effect of having had a study-test trial on each task alone. The mean sum of trials 2 and 3 for Condition T-RFD (23.37) is greater than for trials 1 and 2 of Condition RFD (18.53), t = 2.44, $p < .05$. Nevertheless, the fact remains that a study-test cycle on each task alone was not as effective for simultaneous learning as was simul-

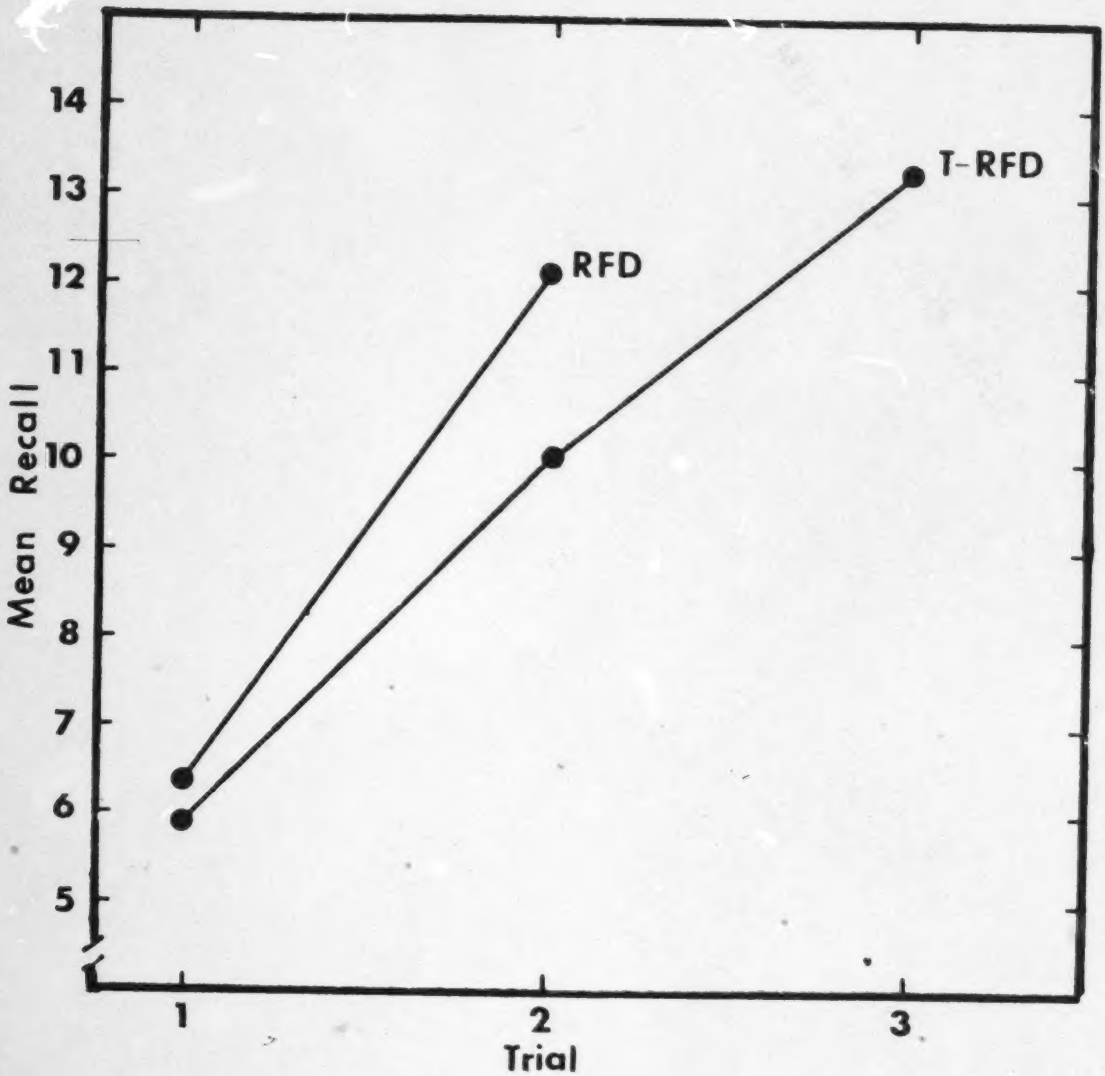

Figure 2. Recall as a function of condition and trial (Experiment 4).

taneous learning per se.

Recognition and frequency judgments. The frequency judgments and the recognition scores (Figure 3) show even more dramatically than did the recall scores the failure of transfer from the initial single trial to simultaneous learning. The mean correlation for the frequency judgments fell from .90 on trial 1 to .82 on trial 2 (t = 6.2?), and the mean recognition errors increased from 2.07 to 6.83 (t = 5.07). The performance on trials 2 and 3 for Condition T-RFD was a little better than on trials 1 and 2 for Condition RFD, but neither for frequency judgments (t = 1.71) nor for recognition (t = .73) were the differences reliable. Thus, although the switch from individual task learning to simultaneous learning did not recult in absolute negative transfer, the positive transfer was minimal. As may be seen in Figure 3, for both tasks there was very little evidence of a latent effect of having had the initial trial. This is shown by the fact that the improvement between trials 2 and 3 for Condition T-RFD was about the same as the improvement between trials 1 and 2 for Condition RFD.

Correlations. For Condition T-RFD, the correlations between the scores on trial 1 and those on trial 2 were .64, .84, and .44 for recall, frequency judgments, and recognition, respectively. The correlations between trials 2 and 3 were identical for all three tasks, .83. The reliabilities were also substantial for Condition RFD, the lowest correlation being .70 for the recognition scores. The interrelationships among the scores on the different tasks were determined by using total scores (all

trials) for each task. For Condition RFD, the values were .10, .33, and .71 for lists 1 x 2, 1 x 3, and 2 x 3, respectively. The corresponding values for Condition T-RFD were .56, .70, and .50. The latter set of correlations must be considered high and thereby do not allow us to make any inferences about differential encoding.

Discussion

Our attempt to establish differential encoding of tasks by using a preliminary trial on each task alone does not appear to have been successful. In fact, the preliminary trial had only a small positive effect on the performance when all tasks were learned simultaneously. It was as if the learning which occurred on the preliminary trial was simply not very appropriate for the learning required when all three tasks were presented simultaneously. We need to ask what is responsible for the lack of positive transfer.

One interpretation might stress the role of forgetting. On the first trial of Condition T-RFD, each task was tested immediately after presenting the items for study. After simultaneous learning, however, the order of the tests was always recall, frequency, and recognition. We do not have exact time differences in the tests for the different tasks but it is conservative to say that the differences in the lengths of the retention intervals would be a few minutes. From unpublished work, we know that forgetting is slight over several minutes following simultaneous learning on study trials, and this is true in spite of the fact that the retention interval may be spent in taking retention tests for other lists. All of this leads to the conclusion that the drops in performance between trials 1 and 2 for recog-

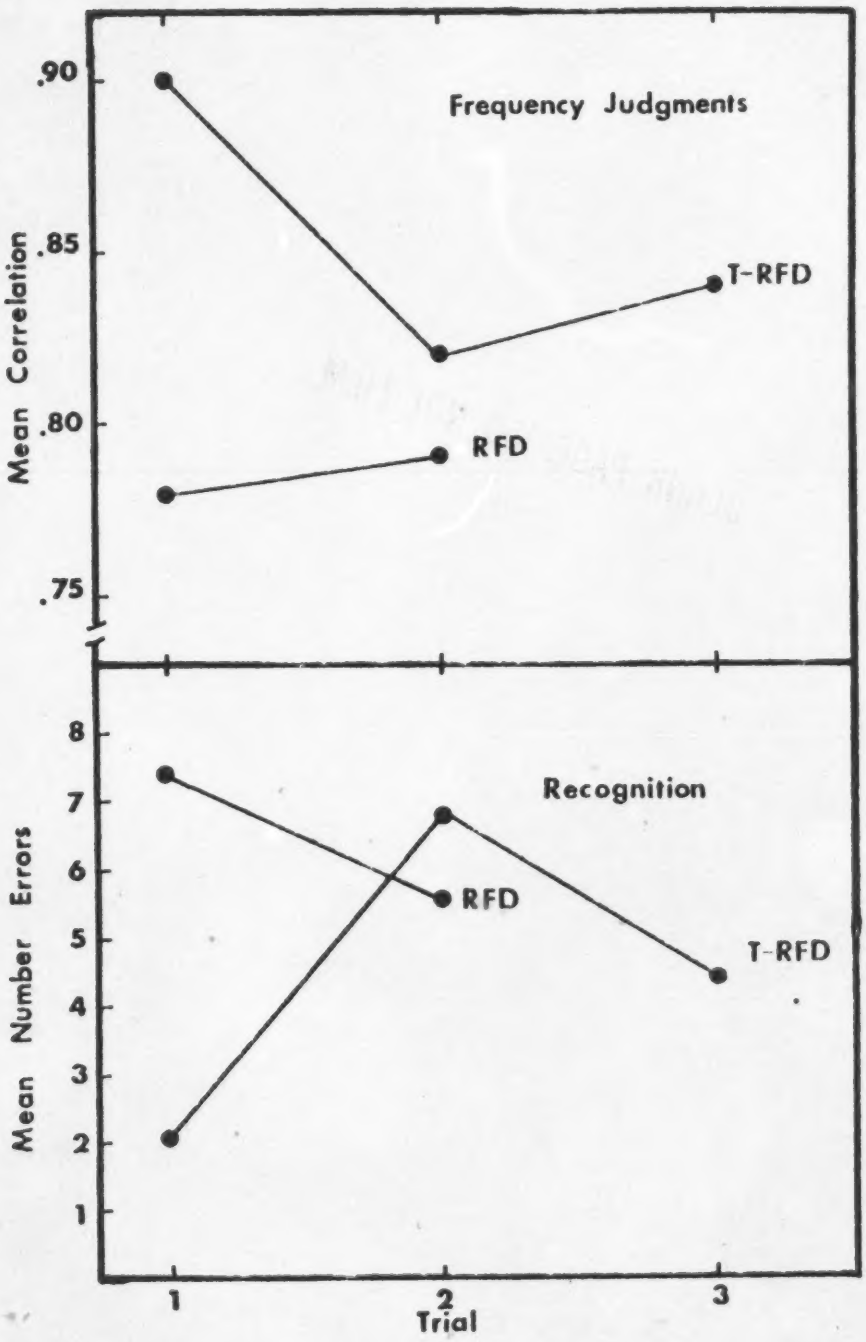

Figure 3. Frequency judgments and recognition as a function of trials and conditions (Experiment 4).

nition and frequency judgments as seen in Figure 3 are probably not due to the longer retention intervals for trial 2 than for trial 1.

The possibility which strikes us as most reasonable would emphasize a change in context. It would seem fairly straightforward to say that the context produced by a series of homogeneous items (single list) was quite different from the context produced when two other quite different lists were learned simultaneously. However, it goes without saying that we had not anticipated this finding at all and the experiment was not analytical with respect to it.

General Discussion

We have been unable to produce evidence supporting the idea that subjects will code lists differently when these lists are being learned simultaneously, and when the nature of the retention test for each list is known. The intertask correlations occasionally suggested differential encoding but there were inconsistencies from experiment to experiment. This is to say that we have not found the correlational evidence to be of much value for the interpretation of the results in the present studies. The critical data were the recall scores on the first list, and these scores gave little indication that differential encoding occurred. It is far more appropriate to draw the opposite conclusion, namely, that all items were studied in the same manner. We cannot, of course, deny other evidence which does suggest that subjects may study differently for a recall and for a recognition test (e.g., Carey & Lockhart, 1973). Perhaps the simultaneous learning situation is not sufficiently sensitive to detect differential encoding. Perhaps the additional assumptions needed to predict differences in perfor-

mance if differential encoding occurs are in error. But the possibility remains that differential encoding of material to fit a particular retention test is not a very likely occurrence.

References

Carey, S. T., & Lockhart, R. S. Encoding differences in recognition and recall. Memory and Cognition, 1973, 1, 297-300.

Garner, W. R. Attention: The processing of multiple sources of information. In E. C. Carterette and M. P. Friedman (Eds.), Handbook of Perception. Vol. 2. New York: Academic Press, 1974.

Kahneman, D. Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Kerr, B. Processing demands during mental operations. Memory and Cognition, 1973, 1, 401-412.

Norman, D. A., & Bobrow, D. G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64.